Image Denoising by Bounded Block Matching and 3D Filtering

Qian Chen and Dapeng Wu

Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32611

Correspondence author: Prof. Dapeng Wu, wu@ece.ufl.edu, http://www.wu.ece.ufl.edu

Abstract

The block-matching with 3D transform domain collaborative filtering (BM3D) achieves very good performance in image denoising. However, BM3D becomes ineffective when an image is heavily contaminated by noise. This is because it allows block-matching to search out of the region where a template block is located, resulting in poor matching. To address this, this paper proposes a bounded BM3D scheme. The novelty of our bounded BM3D is two-folded. First, our scheme partitions an image into multiple regions, and identifies the boundaries between regions. And we restrict block matching search within the region of the template block. Second, to prevent important geometric features such as edges from being removed by collaborative filtering in BM3D, we do partial block matching for different block coherent segments which belong to different regions. Compared to BM3D, the proposed bounded BM3D achieves better visual performance, as well as 0.23dB-1.33dB PSNR increase for heavily noisy images.

Index Terms

BM3D, image denoising, block matching

I. Introduction

The search for efficient image denoising methods still poses a challenge. All the existing methods rely on some explicit or implicit assumptions about the true (noise-free) signal in order to separate it from the noise [1]. In particular, the transform-domain denoising typically assumes that the signal is sparsely represented in the transform domain. However, the great variety of natural images makes it impossible for any fixed 2D transform to achieve good sparsity for all cases. Recently, a novel image denoising strategy based on an enhanced sparse representation in transform domain is introduced in [2], referred to as "BM3D". The enhancement of the sparsity is achieved by grouping similar 2D fragments of the image into 3D data arrays called "groups". Due to the similarity between the grouped blocks, it enables a highly sparse representation in 3D transform domain, so that the noise can be well separated by shrinking the coefficients. To the best of our knowledge, BM3D achieves the best performance for removing additive white Gaussian noise at a reasonable computational cost.

However, if a template block contains geometric objects such as edges, repeating patterns, and texture, BM3D is effective only when the template block can find a large quantity of *good* matches. We notice that BM3D allows block-matching to search out of the region that contains the template block. For a less noisy image, a wide search range may result in crossing different regions that contain different patterns; this increases the chance of finding more good matches. However, when an image is heavily contaminated by noise, true image features are usually concealed in noise, resulting in poor matching. To mitigate this drawback, we propose to partition the image to several regions before block matching. And by restricting the search within the region of a template block, we reduce the risk of finding bad match.

Meanwhile, we would like to point out that although BM3D can generally keep fine image details while removing noise, it blurs edges after collaborative filtering. In 3D transform domain, edges can not be sparsely represented. And the inseparability between noise and edge makes the denoising filter inevitably removes partial energy of edges while mitigating noise. Furthermore, the final stage of BM3D is calculation of the weighted average of the blockwise estimates; however, weighted averaging is a low-pass filter, which blurs edges. In contrast, in our bounded BM3D, if a block contains edge and belongs to more than one region, we conduct partial block matching for each coherent segment rather than the block as a whole. (Here a coherent segment refers to the irregular shaped portion of a block that matches some portion in a template block; e.g., S1 in Fig. 2 is a coherent segment matched to some portion in a template block; S2 in Fig. 2 is another coherent segment matched to some portion in a different template block.) After partial block matching, we apply shape adaptive DCT [3] [4] to each irregular shaped template group. In this way, our method preserves edges and avoids applying DCT transform to edges since DCT is ineffective in representing edges [5].

II. The Bounded BM3D

Similar to BM3D, bounded BM3D follows a two-stage blockwise scheme. In the first stage, we group the matching blocks for every template block in a *noisy* image, and apply hard-thresholding filtering to get a basic estimate. In the second stage, we group the matching blocks for every template block in the *basic* estimate, and apply Wiener filtering to produce the final estimate. Fig.1 illustrates the two stage schemes in a graphic diagram. The difference between BM3D and bounded BM3D is two-folded. 1) After the first stage, the basic estimate is partitioned to regions, and the image boundaries are detected. 2) In the second stage, based on the detected boundaries, we conduct partial block matching for each coherent segment, and restrict search range of block matching within the region of the template block. Then we apply shape adaptive DCT to each coherent segment in Wiener filtering. If a coherent segment happens to be a square block, the procedure is the same except shape adaptive DCT falls back to block DCT. The reason we partition the basic estimate other than the input noisy image is the lack of accuracy in detecting edges in noisy images. Hence partial block matching can only be implemented in Wiener filtering stage where a denoised “cleaner” image - basic estimate is available. We process all overlapping blocks of fixed size in a sliding manner, where process refers to block-matching and filtering in 3D transform domain.

We consider a noisy image $z : \mathbf{X} \to \mathbb{R}$ of the form

$$z(\mathbf{x}) = y(\mathbf{x}) + \eta(\mathbf{x}), \mathbf{x} \in \mathbf{X} \tag{1}$$

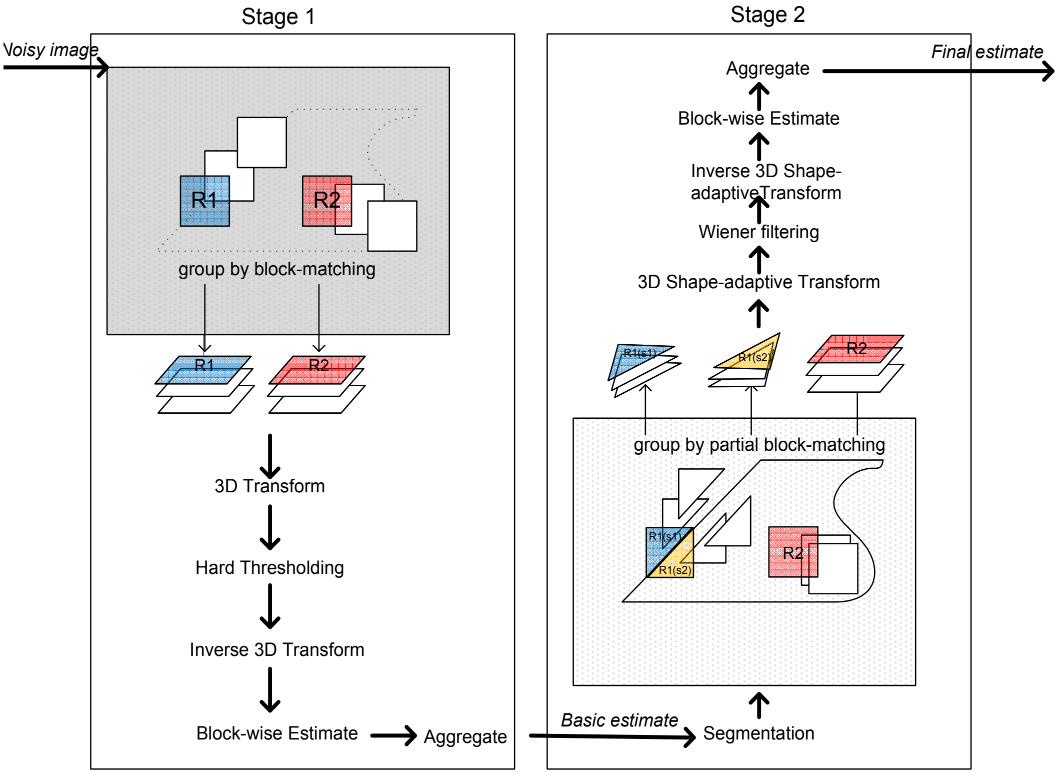


Fig. 1. Diagram of bounded BM3D. Dashed lines in noisy image indicates image boundary is hidden in noise, and it is detected after segmentation in basic estimate.

where $\mathbf{x}$ is a 2D spatial coordinate that belongs to the image domain $\mathbf{X} \subset \mathbb{Z}^2$, y is the true image, and η is i.i.d. zero-mean Gaussian noise with variance σ^2.

With $Z_{\mathbf{x}}$ we denote a fixed size block $N_1 \times N_1$ extracted from z, where $\mathbf{x}$ is the coordinate of the top-left corner of the block. Alternatively, we say that $Z_{\mathbf{x}}$ is located at $\mathbf{x}$ in z. A group of collected 2D blocks is denoted by a bold-face capital letter with a subscript that is the set of its grouped blocks' coordinates, e.g. $\mathbf{Z}_s$ is a 3D array composed of blocks $Z_{\mathbf{x}}$ located at $\mathbf{x} \in S \subseteq \mathbf{X}$. In order to distinguish between parameters used in the first and in the second stage, we respectively use the superscripts "ht" (hard-thresholding) and "wie" (Wiener filtering). We will present the steps of the proposed bounded BM3D in detail in the following parts.

A. *Hard Thresholding Estimate*

We employ hard-thresholding filtering to process the input noisy image to get the basic estimate. This stage follows exactly the same procedures with those in BM3D.

For a reference block located at $\mathbf{x}_R$, $Z_{\mathbf{x}_R}$, its matching blocks group to a stack,

$$S^{ht}_{\mathbf{x}_R} = \{\mathbf{x} \in \mathbf{X} : d(Z_{\mathbf{x}_R}, Z_{\mathbf{x}}) \leq \tau^{ht}\} \tag{2}$$

$d(\mathbf{Z}_{\mathbf{x}_R}, \mathbf{Z}_{\mathbf{x}})$ measures the block distance between the template block and the candidate search block in noisy image, which employs a hard-thresholding prefiltering [2], and τ^{ht} is the distance threshold within which two blocks match. On the other hand, if no match is found, which implies $S^{ht}_{\mathbf{x}_R}$ is null, the method falls back to a 2D problem.

In general, the block-wise estimates can overlap [2], and each block may have more than one estimate. Hence, the basic estimate for $Z_{\mathbf{x}_R}$, denoted as $\hat{Y}^{ht}_{\mathbf{x}_R}$, is a weighted aggregation of all the block estimates located at $\mathbf{x}_R$.

B. Image Boundary Detection

Once we get the basic estimate, we can detect the boundaries by partitioning the basic estimate. The partition method used is left open in our scheme. Experiments show that the selection of partition approach does not have a great impact on the performance gain demonstrated in bounded BM3D. Consequently, in this paper, we employ computational efficient k-means clustering method to partition the basic estimate and get the boundaries thereafter.

C. Wiener Filtering with Image Boundary

Provided that an estimate of the true image is available (e.g. the basic estimate from the hard-thresholding), we can construct an empirical Wiener filter as a natural extension of the thresholding technique [2]. Since the major steps of Wiener filtering in bounded BM3D follows the same procedure as in BM3D, we only give the fundamental modifications required and omit the repetition of the concept. Let us denote the basic estimate by $\hat{y}^{ht} : \mathbf{X} \rightarrow \mathbb{R}$. $E_{\mathbf{x}}$ designates a square block of fixed size $N_1 \times N_1$ extracted from $\hat{y}^{ht}$ and located at $\mathbf{x} \in \mathbf{X}$.

1) Block-matching: In this stage, block matching is conducted in the basic estimate $\hat{y}^{ht}$ rather than the noisy image. Since we restrict the block matching search within the same range of template block, we should first check if the template block contains more than one coherent segment. All the blocks fall in two categories: block contains only one segment (coherent segment is the whole square block) and block contains more than one coherent segment. In the first category, we conduct the whole block matching within the region of the reference block, and cross boundary search is forbidden. In the second category, partial block matching is done for each coherent segment with the help of a mask, which is of the exact same shape as the coherent segment, as illustrated in Fig.2. By doing partial block matching for individual coherent segment that does not contain boundaries, we avoid dealing with edges, so that we also avoid the sparse representation problem of edges in 3D transform domain later in Wiener Filtering.

Suppose a block contains two coherent segments, denoted as S_1, S_2, and $m(S_1)$, $m(S_2)$ are masks corresponding to S_1 and S_2 respectively. By multiplying the candidate block with the mask, every match block will be of the same shape as the coherent segment, and a group of S_1, S_2 shaped blocks are collected for 3D transform denoising later. The block distance measure used in block matching for S_i is written as

$$d_{S_i}(E_{\mathbf{x}_R}, E_{\mathbf{x}}) = \frac{\|E_{\mathbf{x}_R} \bigcap m(S_i) - E_{\mathbf{x}} \bigcap m(S_i)\|^2}{\# m(S_i)^2} \tag{3}$$

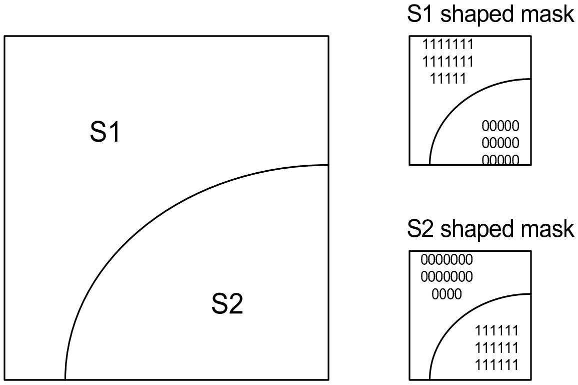


Fig. 2. Mask formation for a block containing two coherent segments S_1 and S_2.

where $E_{\mathbf{x}_R}$ is the template block located at $\mathbf{x}_R$, $E_{\mathbf{x}}$ denotes the candidate matching block located at $\mathbf{x}$, and $\#m(S_i)$ is the number of non-zero elements in $m(S_i)$. Hence the union of $m(S_i)$ with candidate block screens out the elements that are out of the shape of the coherent segment in that candidate block.

Hence, the matched blocks for coherent segment S_i are

$$S^{wie}_{\mathbf{x}_R,i} = \{\mathbf{x} \in \mathbf{X} : d_{S_i}(E_{\mathbf{x}_R}, E_{\mathbf{x}}) < \tau^{wie}\} \tag{4}$$

For expression simplicity, we use s_i to replace $S^{wie}_{\mathbf{x}_R,i}$ and leave out the superscript ”wie“ in the symbols below, where it does not cause any confusion.

With matching block set for S_i, we get two matching groups from basic estimate and noisy image for S_i:

1) $\hat{\mathbf{Y}}^{ht}_{s_i}$ by stacking together the basic estimate fragments $\hat{Y}^{ht}_{s_i}$;

2) $\mathbf{Z}_{s_i}$ by stacking together the noisy image fragments Z_{s_i}.

2) Wiener Filtering: We define the empirical Wiener shrinkage coefficients from the 3D transform coefficients of the initial estimate group as:

$$\mathbf{W}_{s_i} = \frac{|\Gamma_{3D}(\hat{\mathbf{Y}}^{ht}_{s_i})|^2}{|\Gamma_{3D}(\hat{\mathbf{Y}}^{ht}_{s_i})|^2 + \sigma^2}. \tag{5}$$

The collaborative Wiener filtering is applied to the noisy image fragment $\mathbf{Z}_{s_i}$. It is done in the manner of element-wise multiplication of the 3D transform coefficients $\Gamma_{3D}(\mathbf{Z}_{s_i})$ of the noisy data with the Wiener shrinkage coefficients $\mathbf{W}_{s_i}$. Subsequently, the inverse transform $(\Gamma_{3D})^{-1}$ produces the group of estimates

$$\hat{\mathbf{Y}}^{wie}_{s_i} = (\Gamma_{3D})^{-1}(\mathbf{W}_{s_i}\Gamma_{3D}(\mathbf{Z}_{s_i})) \tag{6}$$

located at the matched locations $\mathbf{x} \in s_i$.

To aggregate the overcomplete estimates, for each $\mathbf{x}_R \in \mathbf{X}$, we assign the weight

$$w_{\mathbf{x}_R,i} = \sigma^{-2}\|\mathbf{W}_{s_i}\|^{-2} \tag{7}$$

for the group of estimates $\hat{\mathbf{Y}}^{wie}_{s_i}$.

Let $\hat{Y}^{\mathbf{x}_R}_{\mathbf{x}_m}(x)$ be an estimate of $y^{wie}(x)$, where $\mathbf{x}_m \in S^{wie}_{\mathbf{x}_R,i}$. We extend the arbitrary shaped support of $\hat{Y}^{x_R}_{x_m}(x)$ by zero to simplify the formulation. And the final estimate of the noisy image y^{wie} is the weighted average of all estimates located at the same place, and computed by

$$\hat{y}^{wie}(x) = \frac{\sum_{\mathbf{x}_R}\sum_{\mathbf{x}_m} w_{\mathbf{x}_R,i}\hat{Y}^{\mathbf{x}_R}_{\mathbf{x}_m(\mathbf{x})}}{\sum_{\mathbf{x}_R}\sum_{\mathbf{x}_m} w_{\mathbf{x}_R,i}\chi_{\mathbf{x}_m}(\mathbf{x})} \tag{8}$$

where $\chi_{\mathbf{x}_m} : \mathbf{X} \to \{0, 1\}$ is the characteristic function of the square support of a block located at $\mathbf{x}_m \in \mathbf{X}$.

The problem arises when conducting 3D transform on mask-shaped block group in (5), where the rectangular shaped block transform is often not applicable.

The choice of the 3D transform Γ_{3D} is governed by their energy compaction (sparsity) ability for noise-free image block groups. Experiments in [2] show that 3D transform can be separated to 2D+1D transform, where the well-established decorrelating transforms, such as DCT, DFT, and wavelets can be used for 2D transform, and Haar for 1D transform. In this paper, we choose DCT as 2D transform, and introduce computational attractive shape-adaptive DCT (SA-DCT) to deal with the arbitrary shaped match blocks for each coherent segment.

It is computed by cascaded application of 1D varying-length DCT transforms first on the columns and then on the rows that constitute the considered region [4]. In terms of the basis elements, they are defined as:

$$\Psi^{1D-DCT}_{L,m}(n) = c_m cos(\frac{\pi(2n+1)}{2L}), m, n = 0, \cdots L-1. \tag{9}$$

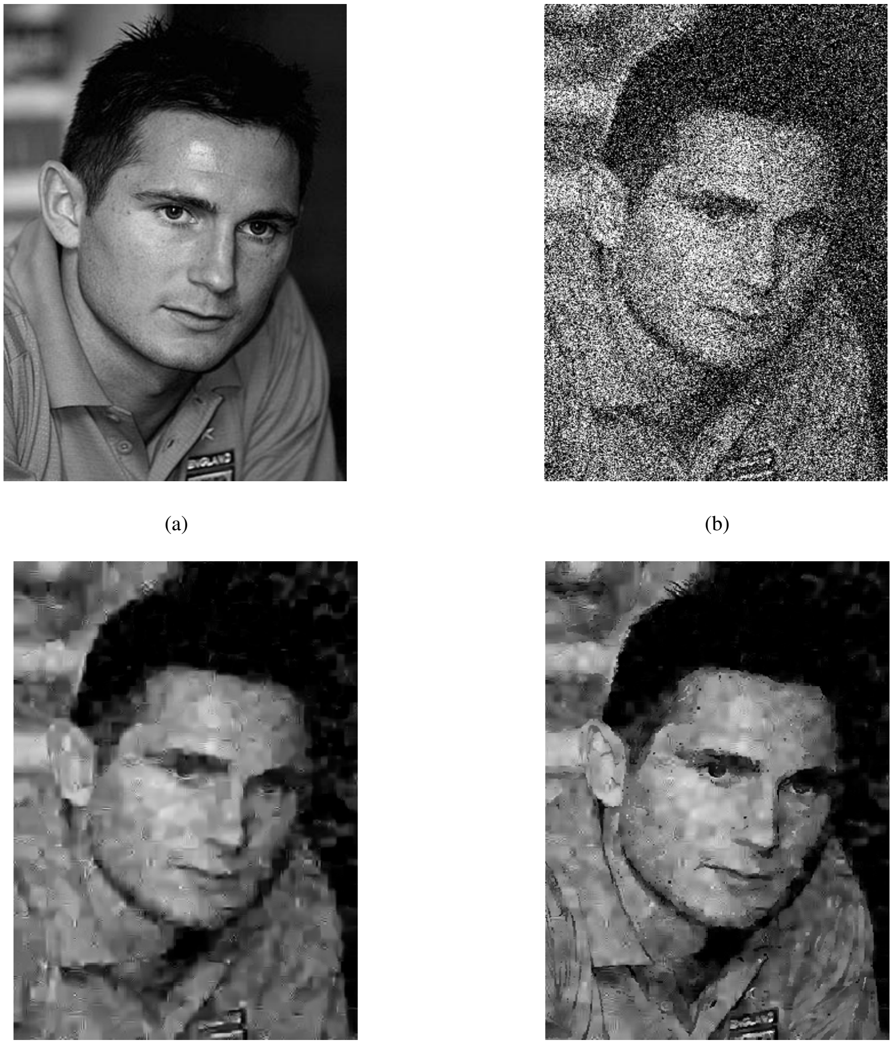
$$c_0 = \sqrt{1/L}, c_m = \sqrt{2/L}, m > 0 \tag{10}$$

Here L stands for the length of the column or row to be transformed. And SA-DCT will degenerate to rectangular 2D-DCT transform if the support is rectangular shaped.

Note that the template block that contains only one coherent segment is a special case of the computation above, with $S_i, i = 1$ only. The mask is simplified to square shape, and hence rectangular shaped block matching and block 2D-DCT is conducted.

III. Experimental Results

We have stated in Section II the limitation of BM3D in highly noise-contaminated image and why bounded BM3D can improve the denoising performance in this case. Therefore, the experiments in this section focus on heavily noisy images. Table I compares the PSNR (dB) performance of BM3D with bounded BM3D in video frames *foreman* and *trevor*, with noise level at $\sigma = 50$ and $\sigma = 100$, respectively. The proposed bounded BM3D achieves 0.23dB-1.18dB gain in these two cases. In Fig. 3, we present the error-free image, heavily noisy image ($\sigma = 100$) and the denoise results using BM3D and bounded BM3D respectively. Similarly, bounded BM3D improves the PSNR by 0.52dB. More remarkably, it shows much improvement in the visual performance, where geometric objects such as edges and texture are more visible to human eyes. We also apply the two schemes in standard test images,*Barbara*(512x512),*Lena*(512x512) and *House*(256x256), to further validate the improvement of the proposed bounded BM3D over BM3D. Fig.4 gives the enlarged fragments of the denoised images, as well



(a)

(b)

(c)

(d)

Fig. 3. Denoising performance comparison between BM3D and Bounded BM3D with heavy noise. (a) Noise-free image. (b) Noisy image with $\sigma = 100$. (c) Denoised image using BM3D,PSNR=23.54dB. (d) Denoised image using bounded BM3D, PSNR=24.06dB.

as the PSNR with the two schemes. We see that bounded BM3D presents better visual effect in texture (the line pattern in Barbara) and object boundary (the boundary in House), as we have expected it achieve. It improves PSNR value in three test images of 0.92dB, 1.30dB and 1.33dB respectively.

To better evaluate the computational complexity of bounded BM3D, we compare the running time (in second) of BM3D and bounded BM3D of the experiment above ($\sigma = 50$). We run the Matlab simulation of the two algorithms on a platform of an Intel Core2 Duo CPU (@2.4GHz and 2.39GHz) PC with MS Windows Vista and 4 GB memory. The results are shown in Table II. It can bee seen that bounded BM3D has a time increase of 15.37% over BM3D on average. The complexity increase is mainly due to the additional step of image boundary detection.

TABLE I

PSNR(dB) comparison between bounded BM3D and BM3D for video frames with heavy noise

	$\sigma = 50$			$\sigma = 100$		
	PSNR (dB)		PSNR increased by	PSNR (dB)		PSNR increased by
	BM3D	Bounded BM3D	Bounded BM3D	BM3D	Bounded BM3D	Bounded BM3D
foreman	25.68	26.23	0.55	21.85	23.03	1.18
trevor	25.30	25.53	0.23	21.66	22.64	0.98

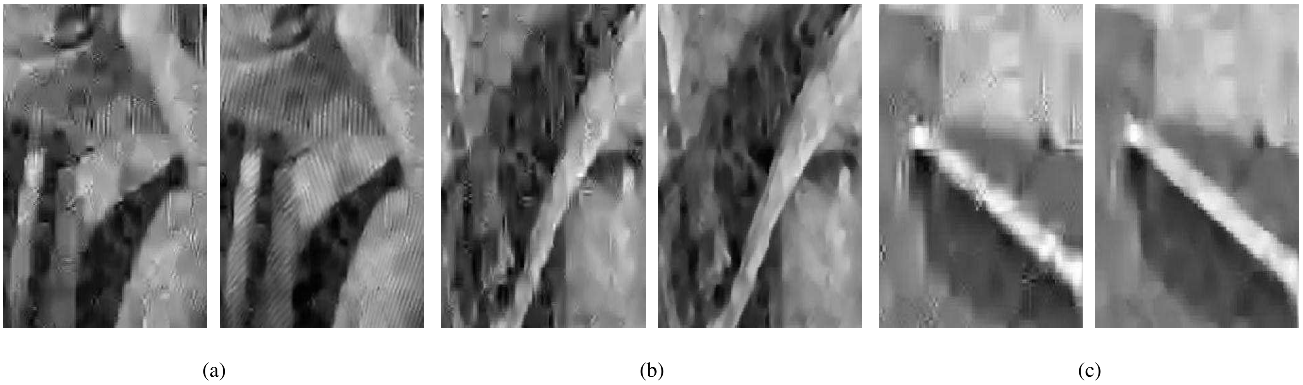
TABLE II

Running Time comparison between bounded BM3D and BM3D for video frames with heavy noise

	Running Time (sec)		Time increased by
	BM3D	Bounded BM3D	Bounded BM3D
foreman	0.40	0.48	18.96%
trevor	0.42	0.47	11.77%

IV. Conclusion

In this paper, we proposed a bounded block-matching with 3D transform domain collaborative filtering method to effectively mitigate noise from images that are highly noisy. The novelty of our bounded BM3D is two-folded. First, after the hard-thresholding filtering, our scheme partitions the basic estimate into multiple regions, and identifies the boundaries between regions. And we restrict block matching search within the region of the template block.



(a)

(b)

(c)

Fig. 4. Fragments of denoised images with BM3D and Bounded BM3D in heavy noise $\sigma = 100$. The left result is by BM3D, the right is by bounded BM3D. (a)Barbara, PSNR=21.44dB vs.PSNR=22.36dB. (b) Lena, PSNR=23.76dB vs.PSNR=25.06dB. (c) House,PSNR=23.53dB vs.PSNR=24.86dB.

Second, to prevent important geometric features such as edges from being removed by collaborative filtering in BM3D, we do partial block matching for different coherent segments which belong to different regions in Wiener Filtering extension. Compared with BM3D, the proposed bounded BM3D achieves better subjective performance, as well as 0.23dB-1.33dB PSNR increase for heavily noisy images.

REFERENCES

[1] A. Buades, B. Coll, and J. Morel, "A review of image denoising algorithms, with a new one," *Multiscale Modeling and Simulation*, vol. 4, no. 2, pp. 490–530, 2006.

[2] K. Dabov, A. Foi, V. Katkovnik, K. Egiazarian, *et al.*, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Transactions on image processing*, vol. 16, no. 8, pp. 2080–2095, 2007.

[3] T. Sikora, "Low complexity shape-adaptive DCT for coding of arbitrarily shaped image segments," *Signal processing. Image communication*, vol. 7, no. 4-6, pp. 381–395, 1995.

[4] A. Foi, V. Katkovnik, and K. Egiazarian, "Pointwise shape-adaptive DCT for high-quality denoising and deblocking of grayscale and color images," *IEEE Transactions on Image Processing*, vol. 16, no. 5, p. 1395, 2007.

[5] M. Yuen and H. Wu, "A survey of hybrid MC/DPCM/DCT video coding distortions," *Signal Processing*, vol. 70, no. 3, pp. 247–278, 1998.